



유전자네트워크 이용한 복잡질환 조절유전자 예측방법 개발

필자는 이미 본지의 지난 호(2010년 3월호)에서 식물유전자네트워크 개발을 소개하면서 유전자네트워크에 대한 기본개념과 향후 농업연구에서의 가능성 등을 소개하였다. 사실 유전자네트워크는 의생명 연구 분야에서 더 오래전부터 관심을 가져왔다. 필자 또한 식물과 더불어 동물의 유전자네트워크의 개발에도 많은 노력을 기울여 이미 2008년도에 기초의학연구에서 매우 유용한 동물모델 중 하나인 꼬마선충의 유전자네트워크를 발표한 바 있다(네이처 제네틱스 40 : 181). 원래 유전자네트워크 개발 당시 기대되었던 주요 응용분야는 새로운 유전자 기능들의 발굴이었다. 그런데 이 연구의 수행 중 본인을 비롯한 공동연구자들은 유전자네트워크를 이용하여 복잡질환의 조절유전자를 발굴할 수 있는 가능성을 보게 되었다. 그래서 우리는 이러한 관찰을 대량의 데이터들을 통해 분석하고 실험을 통해 검증하여 유전자

네트워크를 이용한 복잡질환 조절유전자 예측방법을 제시하였다(개놈 리서치 2010년 6월호 인터넷판).

필자는 본 지면을 통해 복잡질환연구의 중요성과 현재 연구방법들의 한계, 그리고 유전자네트워크가 어떻게 새로운 해결방법을 제시하는지를 소개하고자 한다. 유전자네트워크에 대한 소개는 지난 호에 이미 소개되었으므로 이번에는 생략하기로 한다. 혹시 유전자네트워크가 생소하신 독자들은 2010년 3월호에 게재된 글(네트워크를 이용한 새롭고 효율적인 유전학 연구 기법 제시)을 참고하기 바란다.

대부분의 질환은 수백 개의 유전자들이 복합 작용

인간을 괴롭히는 거의 모든 질환들은 유전적인 요인들이 작용한다. 대학 생화학 교과서에서 많이 등장하는 겸상적혈구빈혈증과 같은 병은 단순히 하나의 유전자 이상으로 질환이 시작되는 전형적인 멘델



글 이인석 연세대학교
생명공학과 교수
insuklee@yonsei.
ac.kr

글쓴이는 연세대학교 생명학과 졸업 후 일리노이주립대에서 생물학 석사학위를, 텍사스주립대에서 미생물학 박사학위를 받았으며, 동대학에서 생명정보학, 시스템생물학분야에서 박사후 과정을 마쳤다.

리언 질환이다. 하지만 이렇게 유전자 한들에 의해 설명될 수 있는 인간의 질환은 현재까지 알려진 질환들 중 5% 미만인 것으로 추측되고 있다. 사실 우리를 위협하고 있는 대부분의 질환들은 수십에서 수백 개의 유전자들의 상태가 복합적으로 작용하여 질환의 발현을 조절하는 복잡질환이다. 예를 들면 많은 사람들의 생명을 위협하는 질환 중의 하나인 암은 현재 약 300~600개의 관련 유전자들이 있을 것으로 추측되고 있다. 이 밖에도 당뇨병, 비만과 같은 대사질환들도 대표적인 복잡질환이다.

수만 명 혹은 수십만 명 중 한 명에게 발생하는 희귀질환인 멘델리언 질환과는 달리 복잡질환들은 우리 주위에서 고통 받는 환자들을 쉽게 찾아볼 수 있을 만큼 보편적인 질환이어서 보편질환이라고도 불리고 있다. 그러므로 인류의 건강증진을 위해서는 복잡질환을 통제할 수 있는 능력이 절실히 필요하다.

복잡질환유전학 연구 어렵게 하는 ‘유전자상호관계’

병을 치료하기 위해서는 병의 원인을 먼저 알아내는 것이 필요하다. 그러므로 유전자의 이상이 질환과 관련된 경우 질환에 영향을 주는 유전자들을 먼저 밝혀내는 것이 가장 급선무일 것이다. 만약 이러한 원인 유전자들을 찾아낸다면 그 유전자들을 타깃으로 하는 신약이나 유전자치료를 통하여 해당 질환을 통제할 수 있을 것이다. 멘델리언 질환의 경우 이러한 원인 유전자의 발굴이 비교적 빠른 시간 내에 이루어질 수 있다. 유전체학의 발전 전에는 최소 몇 년이 걸리던 질환 원인유전자 발굴은 인간유전체의 염기서열 규명과 유전자마커들의 발달 덕분에 현재는 수개월 내에 이루어질 수 있을 정도로 발달되었다. 하지만 복잡질환의 경우는 아직까지도 원인 유전자들의 발굴이 그리 수월하지는 않은 것이 현실이다. 그 주요 이유로는 관련 유전자의 수가 많아 모든 관련 유전자의 발굴에 많은 시간이 요구되고, 각각의 관련 유전자가 질환의 형질에 미치는 영향이 적기 때문에 실험적인 관찰이 어렵기 때문이다.

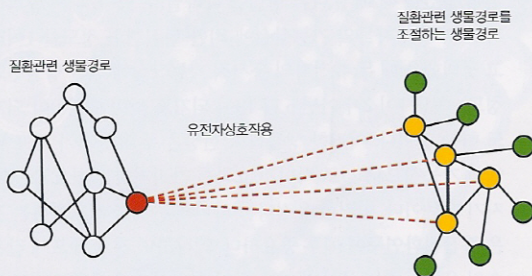
사실 약 5년 전부터 매우 활발하게 진행되고 있는 GWAS라는 고속유전자타이핑과 통계학을 이용한 강력한 연구방법으로 인하여 유전학자들은 복잡질환들에 관련된 유전자들을 밀지않아 대부분 발굴하여 치료법 개발의 근간을 마련할 수 있을 것이라는 희망에 가득 차 있었다. 하지만 최근까지 천문학적인 연구비를 투자하여 얻어낸 결과는 이러한 당초의 기대에는 훨씬 미치지 못하는 것이었다. 복잡질환에 관련된 유전적 요인의 연구는 당초 예상했던 것보다도 훨씬 어려운 것이라는 것이 최근 학계의 공동된 의견이다.

이렇게 복잡질환유전학 연구를 어렵게 하는 여러 가지 이유 중 하나가 바로 유전자상호관계이다. 많은 복잡질환 관련 유전자들은 각각 단독으로 이상이 생겼을 때와 서로 복합적으로 이상이 생겼을 때 매우 다른 형질을 나타낼 수 있다. 다시 설명하자면 각 유전자들의 기능상실에 의한 형질변화 정도를 1이라고 정의했을 때 두 유전자의 동시적 기능상실에 의한 형질변화 정도, 즉 1+1이 2가 아닌 5 나 10이 되기도 하며 때론 0이 되기도 하는 것이다. 이러한 생물학적 현상은 두 유전자가 세포 내에서 기능적으로 서로 시너지 효과를 가지거나 길항효과를 가지기 때문이다. 이러한 유전자들 사이의 관계를 지도화하는 것은 복잡질환연구에 매우 중요하다. 왜냐하면 전체 세포에 같은 수의 유전자에 이상이 생겨도 이러한 시너지 효과를 나타내는 유전자조합의 경우(1+1은 5나 10인 경우)와 아닌 경우(1+1은 2인 경우)의 질환의 정도는 매우 다를 수 있기 때문이다. 또한 복잡질환을 일으킨 한 유전자의 돌연변이를 기능적으로 억제할 수 있는 또 다른 유전자(조절유전자)의 돌연변이가 발견되는 경우(1+1은 0인 경우)에는 이 억제효과를 나타내는 유전자의 발현을 조절함으로써 질환을 치료할 수 있는 가능성도 있는 것이다. 그러므로 이러한 질환관련 유전자들 사이의 유전자상호관계의 발굴은 개별적인 질환관련 유전자의 발굴만큼 복잡질환 연구에 중요한 문제로 대두되고 있다.

두 유전자들 간의 유전자상호작용은 유전자들의 개별적인 기능상실에 의한 형질변화의 정도들과 두유전자의 동시적 기능상실에 의한 형질변화의 정도의 차이를 비교함으로써 비교적 쉽게 발굴될 수 있다. 즉 1+1이 2보다 크거나 작으면 유전적 상호관계가 있다고 말할 수 있다. 하지만 이를 전체 유전체 수준의 문제로 확대하면 이야기가 달라진다. 예를 들면, 인간의 전체 유전자의 수는 약 2만 개로 추정되고 있다. 그러므로 이들 2만 개 유전자들 간의 상호관계의 경우의 수는 2억 개($20,000 \times 20,000 \div 2 = 200,000,000$) 가량이 된다.

수십 개 혹은 수백 개의 상관관계를 실험적으로 테스트하는 것은 비교적 수월한 연구가 될 수 있으나 2억 개를 모두 테스트하는 것은 현재의 기술로는 거의 불가능할 뿐만 아니라 실용적이지 못한 연구방법이다. 참고로 단세포 미생물인 효모의 유전자상호관계에 대한 유전체 수준의 테스트가 캐나다의 토론토 대학에서 현재까지 약 10년 이상 진행되어왔다. 매우 자동화된 시스템을 사용함에도 불구하고 현재까지 테스트된 유전자상호관계의 수는 약 800만 개 정도이다. 그러므로 2억 개의 상호관계를 실험적으로 테스트한다는 것은 훨씬 어렵고 비용이 많이 드는 인간세포주를 이용한 연구에서 무작위적으로 실행한

유전자네트워크를 이용한 질환유전자의 조절유전자(유전자상호작용 관계에 있는 유전자)의 예측법을 도식화한 그림



◀ 질환유전자(적색)는 질환관련 생물경로에 포함되는 유전자이고 유전자상호작용에 의해 알려진 조절유전자(황색)들은 질환관련 생물경로를 조절하는 또 다른 생물경로의 구성 유전자들이다. 이들 알려진 조절유전자의 네트워크상의 이웃유전자(녹색)들이 동일한 생물경로에 포함되어 있다면 이들 역시 동일한 질환유전자를 조절하는 유전자들일 가능성이 높을 것이다. (Genome Research doi:10.1016/j.gr.102743.108의 그림 발췌 수정.)

다는 것은 분명 그리 현명한 계획이 아닌 것은 분명하다.

유전자네트워크를 이용한 복잡질환 연구방법 제시

그렇다면 이를 위한 보다 효과적인 연구방법은 없을까? 필자의 연구팀은 최근 각광받고 있는 시스템생물학 연구방법을 이용하고자 한다. 생명체를 전체 계로 연구하는 시스템생물학에서는 예측모델이 자주 사용된다. 예측모델들은 주로 여러 가지 다양한 오믹스 데이터를 분석하고 통합하여 구축될 수 있으며 필자가 속한 공동연구팀이 주로 이용하는 예측모델이 바로 유전자네트워크이다. 우리의 목표는 유전자네트워크를 이용하여 유전자상호관계 가능성이 높은 유전자 쌍들을 예측하여 최소의 비용으로 최대수의 유전자상호관계를 발굴하는 것이다.

연구팀은 기존에 발굴된 유전자상호관계들이 생물경로들 사이에 많이 존재한다는 기존의 관찰결과에서 아이디어를 시작하였다(그림 참조). 생물경로는 쉽게 독립적인 생리현상을 수행하는 기능단위라고 생각하면 된다. 생리현상을 수행하는 최소 기능단위는 개별 유전자가 아닌 기능적으로 연관된 유전자들의 기능적 집합체인 생물경로이다. 연구팀이 개발한 유전자네트워크는 이러한 생물경로들을 잘 표현하고 있다. 우선 기존에 알려진 질환관련 유전자(적색)와 유전자상호작용관계(적색 점선)에 있는 조절유전자(황색)들을 조사해본 결과 이들 동일한 질환유전자에 상호작용하는 조절유전자들이 우리의 유전자네트워크에서 서로 잘 연결되어있음을 관찰할 수 있었다. 이것은 첫째, 동일 질환유전자와 상호작용하는 조절유전자들이 동일한 생물경로를 이루는 유전자일 확률이 높다는 것을 암시하며, 둘째, 현재 알려진 동일 질환유전자에 대한 조절유전자들의 유전자네트워크 상 이웃 유전자(녹색)들도 또한 조절유

전자일 가능성이 높을 것이라는 가설을 제안하고 있다. 그래서 연구팀은 꼬마선충을 이용해 이 가설을 테스트해보기로 했다.

연구팀이 꼬마선충을 선택한 이유는 크게 세 가지이다. 첫째, 꼬마선충을 이용하면 손쉽게 RNA 간섭을 이용하여 후보 유전자상호작용을 테스트해볼 수 있다. 둘째, 꼬마선충은 인간의 질환관련 유전자와 유사한 유전자들을 많이 가지고 있고, 실제 암, 당뇨와 같은 인간 질환연구에 모델 생명체로 많이 사용되고 있다. 셋째, 연구팀은 이미 세계 최고 수준의 꼬마선충 유전자네트워크를 보유하고 있다. 이에 인간질환을 조절하는 신호 전달경로에 관련된 세 유전자의 꼬마선충 내 유사유전자들에 이미 알려진 조절유전자들의 유전자네트워크 상의 이웃 유전자들을 새로운 후보 조절유전자로 선정한 후, RNA 간섭을 이용, 실험적으로 테스트하였다. 그 결과 유전자네트워크기반의 예측이 기존의 무작위적 예측이나 생물학적 가설기반의 예측보다 평균 7배 이상 효과적임을 관찰하였다.

효과적인 질환조절 유전자 발굴 가능성 높아

그렇다면 위와 같이 유전자네트워크를 이용한 예측을 통한 효과적인 질환조절 유전자 발굴이 과연 가능할 것인가? 필자를 비롯한 대부분의 시스템생물학자들은 매우 가능성이 높다고 예측하고 있다. 하지만 이를 위해서는 우선 예측력이 뛰어난 유전자네트워크가 인간유전체에 대해 필요하고, 꼬마선충에서와 같이 비교적 수월한 실험적 테스트 방법이 또한 필요하다. 많은 기능 유전체학자, 시스템생물학자들이 이를 위해 현재 부단히 연구하고 있으므로 가까운 시일 내에 새로운 가능성을 제시하는 연구결과가 발표될 것을 기대해 본다. **ST**